Interpretation of Shannon Entropies with Various Bases by Means of Multinary Searching Games

Hsuan-Hao Chao[1], **Chang F. Hsu**[1], **Long Hsu**[1] **and Sien Chi**[2]

[1]Department of Electrophysics, National Chiao Tung University, Hsinchu 30010, Taiwan

[2] Department of Photonics, National Chiao Tung University, Hsinchu 30010, Taiwan

E-mail: `schi@mail.nctu.edu.tw`

June 2019

Abstract. Ben-Naim used twenty question games to illustrate Shannon entropy with base 2 as a measure of the amount of information in terms of the minimum average number of binary questions. We found that Shannon entropy with base 2 equal to the minimum average number of binary questions is only valid under a special condition. The special condition is referred to as the equiprobability condition, which requires that the outcomes of every question have equal probability, thus restricting the probability distribution. This requirement is proven for a ternary game and a proposed multinary game as well. The proposed multinary game finds a coin hidden in one of several boxes by using a multiple pan balance. We have shown that the minimum average number of weighing measurements by using the multiple pan balance can be directly obtained by using Shannon entropy with base b under the equiprobability condition. Therefore, Shannon entropy with base b can be interpreted as the minimum average number of weighing measurements by using the multiple pan balance when the multiple outcomes have equal probability every time.

Keywords: Shannon entropy, information, multinary searching game, equiprobability condition

1. Introduction

Shannon entropy, as a measure of the amount of information [1], is an abstract idea that beginning students may feel hard to understand, especially concepts of entropy and information. There are many pedagogical discussions on the meanings of entropy [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12] and information [5, 7, 10, 12, 13, 14], as well as the relation between them [16, 4, 5, 7, 10, 12, 15, 17, 18, 19]. Among these discussions, Ben-Naim used the twenty question (20Q) game to illustrate Shannon entropy in terms of the minimum average number of binary questions [10, 12, 20, 21]. The more questions needed to be asked means the larger amount of information in the game. This game provided a clear picture for students to grasp the idea of Shannon entropy with base 2, but few accessible interpretations for base 3 or more had been proposed in pedagogical materials. Therefore, this paper aims to introduce the concept of Shannon entropy with an arbitrary integer base through games or experience of life, so that students without prior knowledge can understand easily.

Ben-Naim used the 20Q game to interpret Shannon entropy with base 2, but he did not specify the condition that Shannon entropy with base 2 equals to the minimum average number of binary questions. The 20Q game has numerous varieties, and one of them requires finding a coin hidden in one of several boxes with a probability distribution by asking binary (yes-no) questions. The optimal strategy that minimizes the average number of questions divides the boxes into two groups with those probabilities as close as possible and then asks yes-no questions, successively. This is a binary searching process so the 20Q game is a binary game. Ben-Naim claimed that one can prove mathematically that the minimum average number of questions you need to ask to obtain the information is given by the Shannon formula $H = -\sum_{i=1}^{N} p_i \log_2 p_i$, where p_i is the probability of event i, which in our case is the probability that a coin is hidden in box i [22]. However, this statement is only valid under a special condition that was not mentioned.

To illustrate Shannon entropy with base 3, Renyi's fake coin problem [23, 24] can be used as a good start because it is a ternary searching process. Renyi's fake coin problem determines a lighter fake coin among 27 coins by using a double pan balance. Among the 27 coins, any one could be the fake coin with equal probability, which is referred to as the uniform distribution condition. The outcome of the double pan balance is ternary, including the lighter left pan (the fake coin on the left pan), the lighter right pan (the fake coin on the right pan), or balanced (the fake coin in the group outside of the pans). Renyi showed that the minimum average number of weighing measurements is 3 in the optimal strategy. At the first weighing, the coins are divided into three groups with 9 coins in each group. The candidate 9 coins can be determined by weighing any two of the three groups. For the second and third weighing, the coins are divided into three groups as well, with 3 coins and 1 coin in each group. As a result, the fake coin can be determined at the third weighing. However, Renyi did not discuss the relation between this problem and Shannon entropy.

In the following sections, we will show that in the 20Q game, the minimum average number of questions to obtain information equal to Shannon entropy is valid only under a special condition. The special condition is referred to as the equiprobability condition, and its requirement will be specified. For Renyi's fake coin problem, we will show that we can directly use Shannon entropy with base 3 to obtain the minimum average number of weighing measurements to find the fake coin. Renyi's fake coin problem is equivalent to a game of finding a coin hidden in one of several boxes with equal probability by using a double pan balance and can be generalized to the same game without the constraint of a uniform distribution condition. We will also show that Shannon entropy with base 3 of the game is numerically equal to the minimum average number of weighing measurements under the equiprobability condition. Further, if a $(b-1)$-tiple pan balance is available, we will show that we can directly use Shannon entropy with base b to obtain the minimum average number of weighing measurements under the equiprobability condition, where b is an integer greater or equal to 4.

2. Equiprobability Condition for the Binary Game

In the 20Q game, the number of questions you ask each time is one, and the maximum information you obtain each time is one bit when two outcomes have equal probability [25]. Thus, the minimum average number of questions is equal to the amount of information of the game measured by Shannon entropy, as the probabilities of the two outcomes are always equal for all questions, which is referred to as the equiprobability condition. Otherwise, the amount of information gained from one binary question is less than one bit if the two outcomes have unequal probability. To obtain the total information of the game, you need to ask more questions. The minimum average number of questions becomes larger than the Shannon entropy of the game once the probabilities of the two outcomes are unequal. This argument is illustrated mathematically below.

We consider that a coin is hidden in one of five boxes and assume that the optimal strategy is as shown in Figure 1. By following in the same manner as Ben-Naim's calculation [26], the expected amount of information gained from the first question (q_1), the second question by the left strategical pathway (q_{2L}), the second question by the right strategical pathway (q_{2R}), and the third question (q_3) are

$$q_1 = P_T \times H(\frac{P_L}{P_T}, \frac{P_R}{P_T}) \text{ bit}, \tag{1}$$

$$q_{2L} = P_L \times H(\frac{P_a}{P_L}, \frac{P_b}{P_L}) \text{ bit}, \tag{2}$$

$$q_{2R} = P_R \times H(\frac{P_c}{P_R}, \frac{P_D}{P_R}) \text{ bit}, \tag{3}$$

$$q_3 = P_D \times H(\frac{P_d}{P_D}, \frac{P_e}{P_D}) \text{ bit}. \tag{4}$$

The amount of information found by each question is the Shannon entropy with base 2 of the conditional probabilities of the two outcomes weighted by the nodal probability which is the sum of the probabilities of the two outcomes. When the probabilities of the two outcomes are equal, the amount of information of a question is numerically equal to the nodal probability. For example, q_1 is numerically equal to P_T as P_L and P_R are equal. In addition, the sum of q_1, q_{2L}, q_{2R}, and q_3, equals to the total amount of information of this game, which can be evaluated by using Shannon entropy with base 2 along with the entire probability distribution, that is, $q_1 + q_{2L} + q_{2R} + q_3 = H(P_a, P_b, P_c, P_d, P_e)$.

The probability of asking a certain question is exactly the nodal probability. This concept can be used for calculating the minimum average number of questions. The expected number of times of asking the first question (n_1), the second question by the left strategical pathway (n_{2L}), the second question by the right strategical pathway

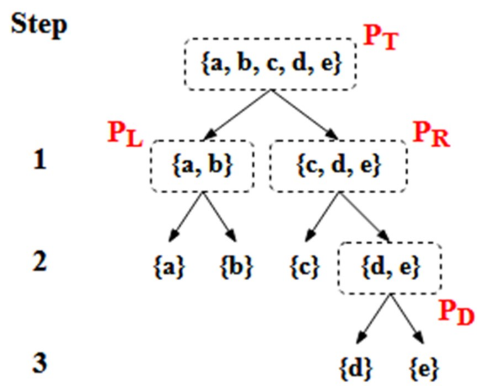

Figure 1. A strategy of asking binary questions in the 20Q game. A coin is hidden in one of five boxes denoted by boxes a, b, c, d, and e with probabilities of P_a, P_b, P_c, P_d, and P_e, respectively. There are four nodes that are marked by the dashed rectangles that are denoted by P_T, P_L, P_R, and P_D, respectively. The nodal probability is the sum of the probabilities of the two outcomes, for example, $P_L = P_a + P_b$.

(n_{2R}), or the third question (n_3) is numerically equal to the nodal probability, P_T, P_L, P_R, or P_D, respectively.

$$n_1 = P_T \times 1 \text{ (question)} = P_T \text{ questions,} \tag{5}$$

$$n_{2L} = P_L \times 1 \text{ (question)} = P_L \text{ questions,} \tag{6}$$

$$n_{2R} = P_R \times 1 \text{ (question)} = P_R \text{ questions,} \tag{7}$$

$$n_3 = P_D \times 1 \text{ (question)} = P_D \text{ questions.} \tag{8}$$

As a result, the value of the minimum average number of questions ($n_1 + n_{2L} + n_{2R} + n_3$) can be obtained by summing these nodal probabilities ($P_T + P_L + P_R + P_D$).

It can be seen that when the two outcomes of a binary question have equal probabilities, the amount of information gained from this question and the expected number of times this question is asked are numerically the same and equal to the nodal probability by comparing q_1, q_{2L}, q_{2R}, and q_3 to n_1, n_{2L}, n_{2R}, and n_3, respectively. In the first question, for example, when P_L and P_R are equal, both q_1 and n_1 become the same and are numerically equal to P_T. Therefore, the value of Shannon entropy with base 2 of the game ($q_1 + q_{2L} + q_{2R} + q_3$) will equal the value of the minimum average number of questions ($n_1 + n_{2L} + n_{2R} + n_3$) under the equiprobability condition. Otherwise, for a question whose two outcomes have unequal probabilities, the amount of information gained from this question becomes the nodal probability times a Shannon entropy less than one and hence will be numerically less than the expected number of times this question is asked. Consequently, the value of Shannon entropy with base 2 of the game will be less than the value of the minimum average number of questions.

In the 20Q game, Shannon entropy with base 2 is numerically equal to the minimum average number of binary questions if and only if under the equiprobability condition which requires that the probabilities of two outcomes are always equal upon all questioning. Thus, in the optimal strategy, both the probabilities of the two groups are designed to be 2^{-1}, 2^{-2}, and 2^{-3} for the first, second, and third questions, respectively, and so on. Accordingly, when a box i is examined at the m_i-th question, the probability that this box i is hiding the coin has to be 2^{-m_i}, where m_i is a positive integer. As a result, the minimum average number of questions becomes $\sum_{i=1}^{N} m_i \times 2^{-m_i}$. For a mathematical proof of the minimum average number of questions as $\sum_{i=1}^{N} m_i \times 2^{-m_i}$, please see the appendix A. Consequently, the value of Shannon entropy with base 2 is equal to the minimum average number of binary questions, mathematically,

$$H = -\sum_{i=1}^{N} 2^{-m_i} \log_2(2^{-m_i}) = \sum_{i=1}^{N} m_i \times 2^{-m_i}. \tag{9}$$

Therefore, we can interpret Shannon entropy with base 2 as the minimum average number of binary questions under the equiprobability condition.

3. Equiprobability Condition for the Ternary Game

In Renyi's fake coin problem, we found that the minimum average weight, three times, can be directly measured by Shannon entropy with base 3 ($H_3 = -\sum_{i=1}^{N} p_i \log_3 p_i$), that is, $-\sum_{i=1}^{27}(1/27)\log_3(1/27) = 3$. Renyi's fake coin problem is equivalent to a ternary game of finding a coin hidden in 27 boxes with uniform distribution by using double pan balance, and we can generalize the game to find a coin hidden in one of several boxes with both a uniform and nonuniform distribution by using double pan balance. We assumed that the weights of the boxes are negligible compared to the coin.

Similar to the 20Q game, it can be deduced that in the ternary game, the number of weighing measurements each time is one, and the maximum information you obtain each time is one trit when the three groups have equal probability according to Shannon entropy with base 3. Hence, the minimum average number of weighing measurements is equal to the amount of information measured by Shannon entropy with base 3 under the equiprobability condition which requires the three outcomes of weighing have equal probability. Otherwise, the amount of information for one weighing is less than one trit. To obtain the total information, you need to weigh more times.

In the ternary game, Shannon entropy with base 3 is numerically equal to the minimum average number of weighing measurements if and only if under the equiprobability condition. Thus, in the optimal strategy, all the probabilities of the three groups are designed to be 3^{-1}, 3^{-2}, and 3^{-3} for the first, second, and third weighing, respectively, and so on. As a result, the probability of box i hiding the coin is restricted to 3^{-m_i}. In addition, box i with probability 3^{-m_i} should be determined at the m_i-th weighing, and thus, the minimum average number of weighing becomes $\sum_{i=1}^{N} m_i \times 3^{-m_i}$. Consequently, Shannon entropy with base 3 is numerically equal to the minimum average number of weighing measurements ,

$$H_3 = -\sum_{i=1}^{N} 3^{-m_i} \log_3(3^{-m_i}) = \sum_{i=1}^{N} m_i \times 3^{-m_i}. \tag{10}$$

Figure 2. A triple pan balance. In the quaternary game, the boxes are divided into four groups. Three of the four groups are separately placed on the three pans to determine the group hiding the coin.

The number of boxes, N, happens to be $3 + (3-1) \times k$ to ensure that the boxes can always be divided into three groups with equal probability every time, where k is an integer greater than or equal to zero.

Therefore, in the equiprobability condition, Shannon entropy with base 3 can be interpreted as the minimum average number of weighing measurements to find the coin hidden in one of several boxes using a double pan balance.

4. Equiprobability Condition for the Multinary Game

Based on the binary and ternary games, we proposed a b-ary game to interpret Shannon entropy with base b. The b-ary game finds a coin hidden in one of several boxes with a probability distribution by using a $(b-1)$-tiple pan balance where b is an integer greater than or equal to 4, for example, a triple pan balance as shown in Figure 2. for the quaternary game. We assumed that the weight of the boxes is negligible compared to the coin. The optimal strategy is to divide the boxes into b groups with the probabilities as close as possible and then weigh any $(b-1)$ groups to successively determine the candidate group. The amount of information in the game is measured by using Shannon entropy with base b, which is defined as $H_b = -\sum_{i=1}^{N} p_i \log_b p_i$.

By following the same rules in the binary game and the ternary game, the requirement of the equiprobability condition for a b-ary game is to restrict the probability of box i hiding the coin to be b^{-m_i}. Additionally, all the probabilities of the b groups are designed to be b^{-1}, b^{-2}, and b^{-3} for the first, second, and third weighing, respectively. The box i with probability b^{-m_i} should also be determined at the m_i-th weighing, and then the minimum average number of weighing measurements becomes $\sum_{i=1}^{N} m_i \times b^{-m_i}$. As a result, the value of Shannon entropy with base b is equal to the minimum average number of weighing measurements by using a $(b-1)$-tiple pan balance,

$$H_b = -\sum_{i=1}^{N} b^{-m_i} \log_b(b^{-m_i}) = \sum_{i=1}^{N} m_i \times b^{-m_i}. \tag{11}$$

The number of boxes, N, happens to be $b + (b-1) \times k$.

Finally, in the equiprobability condition, Shannon entropy with base b can be interpreted as the minimum average number of weighing measurements to find the coin hidden in one of several boxes by using an imaginary $(b-1)$-tiple pan balance.

5. Conclusion

In the binary game, we have found that Shannon entropy with base 2 is equal to the minimum average number of binary questions if and only if under the equiprobability condition. The equiprobability condition in the binary game requires the outcomes of every question have equal probability, leading the probability of box i hiding the coin to be 2^{-m_i}, where m_i is a positive integer. In the ternary game, we introduce Shannon entropy with base 3 to directly evaluate the minimum average number of weighing measurements on the equiprobability condition. The equiprobability condition in the ternary game requires the probability to be in the form of 3^{-m_i} for box i. Further, we proposed a general multinary game that finds a coin hidden in one of several boxes by using a multiple pan balance. Following the same rules, Shannon entropy with base b is equal to the minimum average number of weighing measurements if and only if under the equiprobability condition which restricts the probability of box i hiding the coin to be b^{-m_i} to ensure that the outcomes of weighing have equal probability every time. Therefore, in the equiprobability condition, Shannon entropy with various bases b can be interpreted as the minimum average number of weighing by using a $(b-1)$-tiple pan balance.

6. Appendix

It can be proved that the minimum average number of questions is $\sum_{i=1}^{N} m_i \times 2^{-m_i}$ by exchanging two boxes with different probabilities in the strategy. For example, before the exchange, a box x with probability 2^{-m_x} is determined at the m_x-th question, and a box y with probability 2^{-m_y} is determined at the m_y-th question. After the exchange, box x is determined at the m_y-th question, and box y is determined at the m_x-th question. The average number of questions before the exchange (AN_{bef}) and after the exchange (AN_{aft}) can be written as

$$AN_{bef} = \sum_{i=1}^{N-2} (m_i \times 2^{-m_i}) + m_x \times 2^{-m_x} + m_y \times 2^{-m_y}, \tag{12}$$

$$AN_{aft} = \sum_{i=1}^{N-2} (m_i \times 2^{-m_i}) + m_x \times 2^{-m_y} + m_y \times 2^{-m_x}. \tag{13}$$

The subtraction of AN_{aft} from AN_{bef} always gives a negative number; that is,

$$AN_{bef} - AN_{aft} = (2^{-m_x} - 2^{-m_y}) \times (m_x - m_y) < 0. \tag{14}$$

This result proves that $\sum_{i=1}^{N} m_i \times 2^{-m_i}$ is the minimum average number of questions.

7. Acknowledgments

This research was supported by the Ministry of Science and Technology of the Republic of China grant number MOST 107-2221-E-009-016.

8. Reference

[1] Shannon, C.E. 1948 A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423.
[2] Castans, M. 1962 Entropy and Uncertainty. *Am. J. Phys.* **30**, 521–527.
[3] Hestenes, D. 1970 Entropy and Indistinguishability. *Am. J. Phys.* **38**, 840–845.
[4] Andrew, K. 1984 Entropy. *Am. J. Phys* **52**, 492–496.
[5] Machta, J. 1999 Entropy, information, and computation. *Am. J. Phys* **67**, 1074–1077.
[6] Styer, D.F. 2000 Insight into entropy. *Am. J. Phys.* **68**, 1090–1096.
[7] Brissaud, J.-B. 2005 The meanings of entropy. *Entropy* **7**, 68.
[8] Timberlake, T. 2010 The statistical interpretation of entropy: an activity. *The Physics Teacher* **48**, 516–519.
[9] Swendsen, R.H. 2011 How physicists disagree on the meaning of entropy. *Am. J. Phys.* **79**, 342–348.
[10] Ben-Naim, A. 2012 Shannon's measure of information and the thermodynamic entropy. *AIP Conf Proc.* **1443**, 129–142.
[11] Geller, B.D.; Dreyfus, B. W.; Gouvea, J.; Sawtelle, V.; Turpen, C.; Redish, E. F. 2014 Entropy and spontaneity in an introductory physics course for life science students. *Am. J. Phys.* **82**.
[12] Ben-Naim, A. 2017 Entropy, Shannon's Measure of Information and Boltzmann's H-Theorem. *Entropy* **19**, 48.
[13] Caticha, A. 2007 Information and Entropy. *AIP Conf Proc.* **954**, 11–22.
[14] Losee, R.M. 2017 Information theory for information science: Antecedents, philosophy, and applications. *Education for Information* **33**, 23–35.
[15] Grandy, W.T. 1997 Resource Letter ITP-1: Information Theory in Physics. *Am. J. Phys* **65**, 466–476.
[16] Lemons, D.S. 2013 *A Student's Guide to Entropy*, pp. 143–147 (Cambridge University Press: New York)
[17] Ben-Naim, A. 2011 Entropy: Order or Information. *J. Chem. Educ.* **88**, 594–596.
[18] Leff, H.S. 2012 Removing the Mystery of Entropy and Thermodynamics — Part V. *The Physics Teacher* **50**, 274–276.
[19] Balian, R. 1999 Incomplete descriptions and relevant entropies. *Am. J. Phys.* **67**, 1078–1090.
[20] Ben-Naim, A. 2016 *Entropy, the Truth the Whole Truth and nothing but the Truth*, pp. 32–43 (World Scientific: Singapore)
[21] Ben-Naim, A. 2017 *Information Theory*, pp. 77–94 (World Scientific: Singapore)
[22] Reference 20, p. 37.
[23] Renyi, A. 1987 *A Diary on Information Theory*, pp. 26–27 (Wiley: New York)
[24] Desurvire, E. 2009 *Classical and Quantum Information Theory*, pp. 45–47 (Cambridge University Press: New York)
[25] Reference 21, p. 75.
[26] Reference 21, pp. 81–85.